

# Impact evaluation of child nutrition programmes

Jan Hoorweg

**This article presents a review of current practices and recent developments regarding impact evaluation of nutrition programmes for preschool children in developing countries. Nutrition education, food supplementation and nutrition rehabilitation, the three major components of intervention in child nutrition, generally focus on food consumption and are directed at individual cases. Impact evaluations are generally confronted with situations where different programmes operate in a variety of cultural settings, where the effects of the interventions are generally small, and where methodological difficulties usually weaken the evaluations. Both the interventions and evaluations have to operate under restricting conditions, and different strategies to accommodate this situation are mentioned. It is argued that the contribution of evaluation is not so much in the nature of being a final arbiter, but rather in the nature of helping to seek the optimal combination of local conditions and type of assistance offered.**

The author is research coordinator at the African Studies Centre, PO Box 9507, 2300 RA Leiden, Netherlands.

This article reflects some of the experiences with evaluation of nutrition programmes at the African Studies Centre over the past years. I am grateful to Rudo Niemeyer, Wijnand Klaver, Ted Kliest and Jane Kusin for their comments on an earlier draft of this text.

<sup>1</sup>J.E. Austin and M.F. Zeitlin, eds, *Nutrition Intervention in Developing Countries*, Oelgeschlager, Gunn & Hain, Cambridge, MA, 1981; N.S. Scrimshaw and M.B. Walerstein, eds, *Nutrition Policy Implementation: Issues and Experience*, Plenum continued on page 200

The growing awareness of the nutritional problems in many developing countries, in particular protein-energy malnutrition, has resulted in manifold nutrition interventions.<sup>1</sup> Some interventions, such as agrarian reforms or improvements of marketing systems, are general in nature. Other interventions, such as food-price subsidies and storage loss prevention, aim more directly at increasing food availability to the population. There has also been massive food aid over the past decade. All these interventions have in common that they concentrate primarily on food provision and food availability. Nutrition interventions among preschool children generally focus on food consumption and are usually aimed at individual cases involving direct contact between programme staff and mothers and children. Even so, there exists considerable variation among child nutrition programmes – notably variations in type of approach, nature of the target group and programme objectives.

Because of the substantial resources involved, in the form of food aid, personnel, financial or other assistance, there is a growing demand for evaluation of nutrition programmes. Evaluation has different functions. It serves as a feedback to programme staff about their performance and achievements, and it provides information for national governments and international agencies about the way resources are utilized. It further serves to improve performance of continuing programmes, and to gain insights that can be used in the planning of future programmes. The extent to which each or all of these aims are indeed realized depends on the scope of the evaluation, and is determined by the different needs that agencies may have and the aim to which they want to put an evaluation. Officials and managers have the task of balancing the need for evaluation against the resources that can be set aside and against the degree of interference with day-to-day activities that is acceptable.

Assessing the impact of nutrition intervention is notoriously complicated because of the difficulty of distinguishing between the impact of the programmes and the influence of other factors. Existing evaluation experience is largely with that of child nutrition programmes because of the general priority given to interventions aimed at the well being of children and the considerable resources involved in these programmes. By now, the topic has become something of an independent area of research by itself; witness several recent bibliographies and handbooks.<sup>2</sup>

continued from page 199

Press, New York, 1982; and B.A. Underwood, ed, *Nutrition Intervention Strategies in National Development*, Academic Press, New York, 1983.

<sup>2</sup>A. Burgess, *Evaluation of Nutrition Interventions: An Annotated Bibliography and Review of Methodologies and Results*, 2nd ed, Food and Nutrition Paper No 24, Food and Agriculture Organization, Rome, 1982; I. Figa-Talamanca, *Nutritional Implications of Food Aid: An Annotated Bibliography*, Food and Nutrition Paper No 33, Food and Agriculture Organization, Rome, 1985; R.A. Klein, M.S. Read, H.W. Riecken, J.A. Brown, A. Pradilla and C.H. Daza, eds, *Evaluating the Impact of Nutrition and Health Programs*, Plenum Press, New York, 1979; D.E. Sahn, R. Lockwood and N. Scrimshaw, eds, *Methods for the Evaluation of the Impact of Food and Nutrition Programmes*, United Nations University, Tokyo, 1984; and B. Schurch, ed, *Evaluation of Nutrition Education in Third World Communities*, Hans Huber, Bern, 1983.

<sup>3</sup>G.H. Beaton and J.M. Bengoa, eds, *Nutrition in Preventive Medicine*, World Health Organization, Geneva, 1976; and Austin and Zeitlin, *op cit.* Ref 1.

<sup>4</sup>B. Bosley, 'Nutrition education', in G.H. Beaton and J.M. Bengoa, eds, *Nutrition in Preventive Medicine*, World Health Organization, Geneva, 1976, pp 277-296; and M.F. Zeitlin and C.S. Formacion, 'Nutrition education', in J.E. Austin and M.F. Zeitlin, eds, *Nutrition Intervention in Developing Countries*, Oelgeschlager, Gunn & Hain, Cambridge, MA, 1981, pp 49-72.

<sup>5</sup>H.M. Sinclair and G.R. Howat, *World Nutrition and Nutrition Education*, Oxford University Press, Oxford, 1980.

<sup>6</sup>K.V. Bailey and A. Raba, 'Supplementary feeding programmes', in G.H. Beaton and J.M. Bengoa, eds, *Nutrition in Preventive Medicine*, World Health Organization, Geneva, 1976, pp 297-312; and N.S. Scrimshaw, 'Programs of supplemental feeding and weaning food development', in N.S. Scrimshaw and M.B. Wallerstein, eds, *Nutrition Policy Implementation: Issues and Experience*, Plenum Press, New York, 1982, pp 101-111.

<sup>7</sup>M.A. Anderson, J.E. Austin, J.D. Wray and M.F. Zeitlin, 'Supplementary feeding', in J.E. Austin and M.F. Zeitlin, eds, *Nutrition Intervention in Developing Countries*, Oelgeschlager, Gunn & Hain, Cambridge, MA, 1981, pp 25-48.

<sup>8</sup>R. Cook, 'Is hospital the place for treatment of malnourished children?', *Environmental Child Health*, Vol 17, 1971, pp 15-25; and J.M. Bengoa, 'Nutrition rehabilitation', in G.H. Beaton and J.M. Bengoa, eds, *Nutrition in Preventive Medicine*, World Health Organization, Geneva, 1976, pp 321-334.

<sup>9</sup>M. Beaudry-Darisme and M.C. Latham, 'Nutrition rehabilitation centres: an evaluation of their performance', *Journal of Tropical Pediatrics and Environmental Child Health*, Vol 19, 1973, pp 299-332; and I.D.

continued on page 201

## Child nutrition programmes

Different authors have classified nutrition programmes for young (preschool) children in somewhat different ways.<sup>3</sup> Three basic types of programmes are usually distinguished: nutrition education, food supplementation and nutrition rehabilitation.

Nutrition education aims to improve food-related practices and individual food consumption, particularly that of small children, and varies from mass media and formal education to the instruction of small groups or individuals.<sup>4</sup> Since the instruction often focuses on the introduction of new practices and better utilization of existing household resources, it has to be adapted to existing practices and local food resources. Existing knowledge about the conditions which make for success or failure of nutrition education in the rural areas of developing countries is still limited.<sup>5</sup>

Supplementary feeding consists of the provision of foods, free of charge or at low cost, to vulnerable groups to cover deficiencies in their habitual diet.<sup>6</sup> It was estimated that in 1979 over 50 million children in developing countries received some kind of food supplement.<sup>7</sup> Supplementary feeding can take the form of 'on-site' feeding or 'take-home' programmes. Initially, the main item of food supplementation was milk as an important source of protein. Today, programmes generally try to offer a balance between high-protein and high-calorie foods. There is also a growing trend to utilize local foods instead of imported food commodities, if at all possible.

Nutrition rehabilitation concerns the treatment of severe cases of protein-energy malnutrition. Early experiences with hospital treatment were not positive, mortality rates were high, hospitalization usually lengthy, and the number of relapses tended to be high.<sup>8</sup> This led to alternative treatment at nutrition rehabilitation centres, which have been introduced in many countries.<sup>9</sup> At day-care centres, children are brought daily for several hours, long enough to be given several meals. At residential centres, children and mothers usually stay for a period of several weeks and the mothers are closely involved in feeding their children back to health using locally available foods and indigenous cooking methods.<sup>10</sup>

In time, programmes have tended to borrow from each other, for example, incorporating nutrition instruction in food supplementation, so that programmes now usually consist of a mix of elements. Individual programmes differ considerably as to the form that they may take. Programmes also differ regarding a number of secondary characteristics, such as the nature of the target group (whether individual child or household; whether geographically, economically or age-defined),<sup>11</sup> the exposure of the participants (frequency, intensity and duration of contact)<sup>12</sup> and the nature of accompanying public health measures (immunization, hygiene, family planning).<sup>13</sup>

Child nutrition programmes have to operate under the extreme diversity of circumstances found in many developing countries. These include the profound difference between rural and urban living conditions, and the vast ecological variation, ranging from arid lands to fertile highlands and tropical forests. They also include cultural differences, which vary from the eating habits of pastoral populations with a milk diet to those of agriculturalists dependent on cereals, roots or tubers. The implications of these differences for the functioning and the impact of programmes is usually a matter of guesswork.

continued from page 200

Beghin and F.E. Viteri, 'Nutrition rehabilitation centers: an evaluation of their performance', *Journal of Tropical Pediatrics and Environmental Child Health*, Vol 19, 1973, pp 404-416.

<sup>10</sup>W.A.M. Cutting, 'Nutrition rehabilitation', in D.S. McLaren, ed, *Nutrition in the Community*, Wiley, New York, 1983, pp 321-337.

<sup>11</sup>The actual targeting, however, is not easy and programmes are faced by a tradeoff between the error of inclusion (including cases that are not really in need of the intervention measures) and the error of exclusion (excluding cases that require assistance). Much depends on the prevalence of malnutrition in the population. Where prevalence is high it appears that efforts to discriminate among children and/or families are less effective than among populations where prevalence is low. R.J. Timmons, R.I. Miller and W.D. Drake, *Targeting: A Means to Better Intervention*, Community Systems Foundation, Ann Arbor, MI, 1983; and R. Timmons, R. Miller and W.D. Drake, 'Selecting participants for community nutrition interventions in developing countries', in M.J. Forman, ed, *Nutritional Aspects of Food Aid*, ACC/SCN (FAO), United Nations, Rome, 1985, pp A1-A20.

<sup>12</sup>In general, programmes designed for high frequency of contact are of limited duration; whereas when duration is prolonged the frequency of contact tends to be less. There is often considerable difference between planned exposure and the exposure actually realized. In practice, mothers and children often drop out at an early stage, while the visits during the actual period of contact may be less frequent than planned.

<sup>13</sup>The well known synergism between malnutrition, childhood infections and parasite infestations necessitates preventive measures: immunization and hygiene, especially sanitation and proper handling of drinking water and food. R. Cook, 'Immunization programmes in the context of prevention of malnutrition', in G.H. Beaton and J.M. Bengoa, eds, *Nutrition in Preventive Medicine*, World Health Organization, Geneva, 1976, pp 268-276; and J.E. Austin, T.K. Belding, R. Brooks, R. Cash, J. Fisher, R. Morrow, N. Pielemeier, D. Pyle, J.D. Wray and M.F. Zeitlin, 'Integrated nutrition programs and primary health care', in J.E. Austin and M.F. Zeitlin, eds, *Nutrition Intervention in Developing Countries*, Oelgeschlaeger, Gunn & Hain, Cambridge, MA, 1981, pp 123-136.

The inclusion of family planning in nutrition programmes is more controversial; in some countries attention to family planning may make a programme unacceptable to large sections of the population.

<sup>14</sup>K.M. Hennigan, B.R. Flay and A. Haag, 'Clarification of concepts and terms commonly used in evaluative research', in R.E. Klein, M.S. Read, H.W. Riecken, J.A.

continued on page 202

## Impact evaluation

Given the range of existing interventions and the possible combinations of elements, it is not surprising that the actual implementation and effectiveness of nutrition programmes show equally large variations. This is where the need for evaluation arises. In its broadest sense, evaluation concerns any information about operations and impact of programmes or policies. In this sense it may consist of the fleeting impression of visiting experts as well as hard empirical evidence.<sup>14</sup>

An important distinction is that between process and impact evaluation.<sup>15</sup> This terminology refers to the conceptual difference between the services offered by the programme and the actual effects or impact of the programme. Process evaluation is concerned with implementation, that is, the degree and manner in which the planned services are indeed delivered. Impact evaluation is concerned with effects, that is, the changes that the programme actually achieves. The first kind of evaluation is more simple than the second one and relatively easy to realize; the second is more complex and relies heavily on research methods drawn from the social sciences.

Definitions of impact evaluation generally include reference to the measurement of programme effects by 'objective' and 'systematic' means. The first refers to the use of reliable measuring instruments with standard routines; the second to the comparison of different groups of recipients (and non-recipients) in such a way that it reveals the impact of the intervention, irrespective of other factors that influence child nutrition. Impact evaluation has two major components: the indicators selected to reflect programme impact, and the design used for the comparison of different groups of recipients.

## Indicators

Indicators can consist of very different measuring instruments, provided they reflect a meaningful aspect that the intervention aims to influence.<sup>16</sup> For child nutrition programmes they usually consist of some aspect of nutritional cognition and nutrition behaviour of mothers, or food consumption and nutritional status of children. Generally they may include the following:

- nutritional cognition (nutritional knowledge, nutritional attitudes);
- food practices (food production, food storage, food preparation, food distribution);
- food consumption (dietary practices, food intake);
- nutritional status (anthropometry,<sup>17</sup> clinical examination, biochemical indicators).

Since improvements in nutritional status are usually the desired outcome of intervention and are often regarded the 'final' outcome, many impact evaluations have been limited to assessment of anthropometry. This offers a very restricted form of impact evaluation. In the cases where such programmes are not successful and no improvements in nutritional status are observed, it is often not possible to discover the exact reason for this failure: whether it is due to inadequate intervention methods, poor programme implementation or perhaps to hindering social influences. Moreover, in all programmes, however successful, there is usually a minority of cases that do not make good progress. Often there is an urgent need to know why these children do not

continued from page 201

Brown, A. Pradilla and C.H. Daza, eds, *Evaluating the Impact of Nutrition and Health Programs*, Plenum Press, New York, 1979, pp 387-432.

<sup>15</sup>A similar, but less often used, distinction is that between formative and summative evaluation.

<sup>16</sup>For detailed technical discussions of indicators and reference standards, see Sahn *et al*, *op cit*, Ref 2; and WHO, *Measuring Changes in Nutritional Status: Guidelines for Assessing the Nutritional Impact of Supplementary Feeding Programmes for Vulnerable Groups*, World Health Organization, Geneva, 1983.

<sup>17</sup>Growth monitoring is an important diagnostic tool. More than 200 different growth charts are estimated to be currently in use, in over 80 countries. UNICEF, *The State of the World's Children*, Oxford University Press, London, 1985.

<sup>18</sup>G.H. Beaton, 'Evaluation of nutrition interventions: methodologic considerations', *American Journal of Clinical Nutrition*, Vol 35, 1982, pp 1280-1289.

<sup>19</sup>G.A.O. Alleyne, R.W. Hay, D.I. Picou, J.P. Stanfield and R.G. Whitehead, *Protein-Energy Malnutrition*, Edward Arnold, London, 1977; J.-P. Habicht and W.P. Butz, 'Measurement of health and nutrition effects of large-scale intervention projects', in R.E. Klein, M.S. Read, H.W. Riecken, J.A. Brown, A. Pradilla and C.H. Daza, eds, *Evaluating the Impact of Nutrition and Health Programs*, Plenum Press, New York, 1979, pp 133-170; S. Schofield, *Development and the Problems of Village Nutrition*, Croom Helm, London, 1979; and M. Zeitlin, 'Intervention evaluation', in M.F. Zeitlin and C.S. Formacion, *Nutrition Intervention in Developing Countries. Study 2. Nutrition Education*, Oelgeschlager, Gunn & Hain, Cambridge, MA, 1981.

<sup>20</sup>J. Hoorweg, 'The control of non-treatment variables: necessity or illusion?', in B. Schurch, ed, *Evaluation of Nutrition Education in Third World Communities*, Hans Huber, Bern, 1983, pp 153-165.

<sup>21</sup>Also called confounding variables, nuisance variables, background variables or extraneous variables.

<sup>22</sup>Hennigan *et al*, *op cit*, Ref 14.

<sup>23</sup>D.T. Campbell and J.C. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Rand McNally, Chicago, 1966; and T.D. Cook and D.T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Houghton Mifflin, Boston, 1979.

<sup>24</sup>J. Cohen, 'Multiple regression as a general data-analytic system', in E.L. Struening and M. Guttentag, eds, *Handbook of Evaluation Research*, Sage Publications, Beverly Hills, 1975, pp 570-595; and R.E. Klein, R.J. Townsend, A. Praun and M. Fischer, 'The practice of impact evaluation of nutrition education programs', in B. Schurch, ed, *Evaluation of Nutrition Education in Third World Countries*, Hans Huber, Bern, pp 136-151.

improve, and it is difficult to discover the reason if data collection has been limited to nutritional status. It is therefore often advisable to include measures of other outcomes as well. A distinction is generally drawn between proximal and distal outcomes. In the case of nutrition education, a proximal outcome would be improved nutritional cognition and nutrition behaviour of the mothers. Improvements in nutritional status are more distal and the more distal the outcome the larger the potential array of variables that also influence it. By using different indicators, it is possible to test and compare alternative causal explanations or different causal pathways.<sup>18</sup>

## Design and control of variables

It is usual to distinguish between treatment variables (various interventions); outcome variables (the proximal and distal outcomes) and non-treatment variables (other determinants of nutrition behaviour and nutritional status). Different authors have expounded the many factors that may influence nutrition behaviour and nutritional status.<sup>19</sup> Variables can range from macro-factors such as ecological and cultural differences between communities, to meso-factors that mainly concern differences between households within communities, and micro-factors concerned with variation within households.<sup>20</sup> The major function of research designs is the control of non-treatment variables,<sup>21</sup> that is, the reduction of the influence of variables that fall outside the purpose of the evaluation. Control serves to ensure that other determinants of nutrition behaviour and nutritional status do not offer rival explanations for any purported relation between treatment and outcome, and serves to reduce residual variance to increase the power or sensitivity of the evaluation. Three different approaches to the control of variables exist: experimental control, quasi-experimental designs and statistical control.<sup>22</sup>

Experimental control is realized by means of experimental designs: measures of outcome variables are compared across two or more groups of people who have received different amounts of treatment. These groups are formed by randomization, that is random assignment of persons to groups, so that the latter are equivalent in every respect except treatment. Any differences in outcomes can therefore be imputed to the effect of the treatment.

Quasi-experimental designs also involve comparisons between different treatment groups and control groups not exposed to treatment. Unlike in experimental designs, these groups are not formed by random assignment. Such non-equivalent comparison groups can be selected in different ways.<sup>23</sup>

While experimental and quasi-experimental designs make use of comparison groups, statistical control is characterized by statistical adjustment of non-treatment variables. Statistical control, notably multivariate analysis, makes use of the observed correlations between variables to remove the influence of non-treatment variables on outcome variables.<sup>24</sup>

In general, control over non-treatment variables is best achieved through randomization (ie experimental control). Both statistical control and quasi-experimental designs are flawed in the sense that the possibility always remains that variables not included in the design or analysis contribute to observed results. Statistical control is generally

considered superior to quasi-experimental designs because it gives more opportunity to remove residual variance.<sup>25</sup> Experimental and quasi-experimental studies often use statistical adjustment to control for non-treatment variables that cannot be randomized conveniently or to counter imperfections of design.<sup>26</sup>

### Current practice

In practice most evaluations rely on quasi-experimental designs. A recent, annotated bibliography on nutrition education in Third World communities shows that very few evaluations of nutrition education programmes relied on randomization or statistical control.<sup>27</sup> The reasons for this are fairly obvious.

Experimental control is difficult to realize because in the case of continuing programmes it is, for ethical reasons, difficult to allocate subjects randomly to different groups. A more general problem is to ensure that mothers and children in the different groups are subsequently kept separate. A further complication arises from the fact that it is necessary to study and compare groups before as well as after treatment, so that extended periods of time have to be covered.

Statistical control requires advance knowledge about the determinants of nutrition behaviour and nutritional status in order to know what should be measured; such knowledge for Third World communities is limited. Other complicating factors are that statistical control requires that fairly large numbers of people be studied, and that it necessitates intricate computations for which computer facilities as well as statistically trained personnel and time for analysis are required.

The methodology of quasi-experimental designs has made considerable progress over the past decades, particularly with the growing interest in evaluation and social experimentation. In fact, recent advances are so sophisticated that some quasi-experimental designs are more complex and more costly than true experimental designs.<sup>28</sup> Nevertheless, the essential weakness remains the possible selection of non-equivalent comparison groups. By selecting for comparison groups of people who differ in exposure to treatment but who may also differ in other respects, such as motivation, education, income, age, or health, virtually any determinant of nutrition behaviour or nutritional status can be artificially introduced, confounding the results.

Evaluations of nutrition programmes in developing countries are usually severely limited in their choice of research strategy and have to make methodological concessions. The objectives of programmes are often not well defined and the daily organization of the programmes necessitates practical research solutions which may be less than optimal. On the other hand, the objective of quasi-experimental designs is not so much to guard against every possible source of 'error', but rather to control those sources of 'error' likely to emerge in a given situation.<sup>29</sup> In general, evaluation should be extensive to ensure that the role of the social environment can be taken into proper consideration.

Recent comprehensive reviews covering evaluation studies from various parts of the world note the modest influence that interventions generally have, if any impact occurs. Moreover, recipients usually represent only a small proportion of the total population, which further reduces the social importance of interventions. A review of food aid evaluations tells 'a rather dismal story and nutritional impact has rarely

<sup>25</sup>Statistical control has the advantage that it is more flexible because many research decisions can be left until the stage of data analysis, when the actual, observed relations between relevant variables can be taken into account.

<sup>26</sup>F.N. Kerlinger, *Foundations of Behavioral Research*, 2nd ed, Holt, Rinehart and Winston, London, 1973; and R.A. Cooper and A.J. Weekes, *Data, Models and Statistical Analysis*, Philip Allen, Oxford, 1983.

<sup>27</sup>B. Schurch and L. Wilquin, *Nutrition Education in Communities of the Third World: An Annotated Bibliography*, Nestle Foundation, Lausanne, 1982.

<sup>28</sup>L. Saxe and M. Fine, *Social Experiments: Methods for Design and Evaluation*, Sage Publications, Beverly Hills, 1981.

<sup>29</sup>C.H. Weiss, *Evaluation Research: Methods of Assessing Program Effectiveness*, Prentice Hall, Englewood Cliffs, NJ, 1972.

been substantiated'.<sup>30</sup> Other reviewers of child feeding programmes conclude that 'ongoing food distribution programmes seem to have relatively little impact'.<sup>31</sup> From a recent annotated bibliography on nutrition education in Third World countries it also appears that, where impact was measured, it was usually found to be modest.<sup>32</sup>

The question must be raised whether this reflects a genuine absence of effects or whether it is perhaps a consequence of poor evaluations. Ultimately, there are two kinds of methodological errors: false positive, concluding that effects exist where they do not exist, and false negative, concluding that no effects exist where they do exist. In the latter case, either the study design or the selected indicators are not sufficiently sensitive to measure the perhaps minor effects of the respective interventions. This type of error usually receives less attention, but it has been suggested that it is in fact the more common and could be the reason for the reported absence of effects in many evaluation studies.<sup>33</sup>

Design faults, however, mostly take the form of selection of groups that are spuriously different, that is, groups that will show themselves to be different on outcome measures but for reasons other than the intervention. This type of error would result in a greater number of significant findings. This is clearly not the case and it seems unlikely that researchers have been making mistakes in the opposite direction, by selecting groups that are not sufficiently different in exposure. A series of detailed studies of different programmes in Uganda and Kenya that went beyond the usual research effort in terms of attention to data collection and control of variables, also revealed only modest effects in the case of successful programmes.<sup>34</sup>

The studies that do report effects are, in fact, often discredited as having methodological weaknesses and as leaving room for possible alternative explanations. This is inherent to the very nature of non-experimental methods. The problem is not so much that imperfections exist, but how to cope with them. The intelligent use of findings from different studies can lead to an overall picture which is more convincing than any one particular set of data. In this sense, evaluation is not different from social research in general, where only the existence of mutually supporting but independent data sets eventually leads to accepted theoretical insights.

## Recent developments

The realization of this situation has elicited different reactions. Those concerned with the quality of the evaluations advocate stricter methodologies and more elaborate designs, while those concerned with the implementation of nutrition intervention stress the urgent need for further empirical insights, and often advocate a relaxation of methodological rigour. A practical suggestion in this connection is to look for so-called 'gross' outcomes with the help of quick studies, before trying to establish 'net' outcomes with the help of stricter designs.<sup>35</sup> The danger is that the gross outcomes will start to lead a life of their own, and that the required further evaluation is not undertaken.

In general, there is a strong tendency to strive for a 'crucial' evaluation, in analogy with the 'crucial experiment': the mythical study that will give a final answer once and for all. In addition, the results of the evaluation are often expected to be cross-nationally relevant. In the case of nutrition interventions in Third World countries, this research

<sup>30</sup>D.E. Sahn, 'Methods for evaluating the nutritional impact of food aid projects: lessons from past experience', *Food and Nutrition Bulletin*, Vol 6, 1985, pp 1-14; and D.E. Sahn and R.M. Pestronek, *A Review of Issues in Nutrition Program Evaluation*, Program Evaluation Discussion Paper No 9, USAID, Washington, DC, 1981.

<sup>31</sup>Beaton, *op cit*, Ref 18, p 1284; and G.H. Beaton and H. Ghassemi, 'Supplementary feeding programmes for young children in developing countries', *American Journal of Clinical Nutrition*, Vol 35, 1982, pp 864-916 (Supplement 4).

<sup>32</sup>Schurch and Wilquin, *op cit*, Ref 27.

<sup>33</sup>Sahn, *op cit*, Ref 30, pp 1-14.

<sup>34</sup>J. Hoorweg and I. McDowell, *Evaluation of Nutrition Education in Africa: Community Research in Uganda, 1971-1972*, Mouton, The Hague, 1979; and J. Hoorweg and R. Niemeyer, *Intervention in Child Nutrition: Evaluation Studies in Kenya*, Kegan Paul International, London, 1988.

<sup>35</sup>J.B. Mason and J.P. Habicht, 'Stages in the evaluation of ongoing programmes', in D.E. Sahn, R. Lockwood and N. Scrimshaw, eds, *Methods for the Evaluation of the Impact of Food and Nutrition Programmes*, United Nations University, Tokyo, 1984, pp 26-45.

approach is not realistic. Not only are there many different types of programmes, but similar programmes have to work among very different population groups and often under unique circumstances. Each nutrition programme therefore faces a multitude of conducive or hindering conditions which means that there is no single 'best' intervention, but that certain interventions function more effectively under some conditions. It is on this aspect that attention should focus. Evaluation should not be conceptualized as providing a set of ultimate and objective standards for clearcut decisions to continue or terminate a particular intervention. Nutrition programmes generally exist because there is a 'felt need', either humanitarian or political, and the contribution of evaluation is more in the nature of helping to seek the optimal combination of local conditions and type of assistance.

It has been suggested that in the case of food supplementation, the intervention can be regarded as an income transfer.<sup>36</sup> The economic value of this income can be calculated and regarded as the benefit, without distinguishing whether the food is eaten by a possible target child, eaten in addition to existing food resources, eaten instead of food that the family would otherwise have eaten, or even sold in the market. Although this reasoning has something of a disappearance trick, solving some evaluation problems by a change of definitions, it does point at an area of future research.<sup>37</sup>

Impact evaluations must be prepared to assess very different interventions and to look for small effects against a background of profound social and economic variation. This suggests several possible research strategies, such as improved control over study conditions, the analysis of 'weak' statistical effects and the use of other than quantitative research methods. The analysis of so-called weak statistical effects requires series of similar studies of similar programmes to be undertaken. The many and diverse cultural settings of nutrition programmes also lead to a need for replication of programme evaluations under different conditions. Qualitative research methods, notably of an anthropological nature, provide behavioural insights which may eventually suggest programme improvements, but which generally do not give a quantitative answer as to whether or not the intervention has an effect. Usually, however, a solution to the evaluation dilemma is sought in improved study designs with greater control over non-treatment variables. This is not only difficult to realize, it is nearly always costly and time-consuming.

Many authors conclude that there is general agreement that evaluation should be a routine programme activity, that field staff must be involved at all stages and that any results should rapidly be fed into the programme concerned.<sup>38</sup> It is in this perspective that the calls for 'in-built monitoring and evaluation' must be regarded.<sup>39</sup> These monitoring and evaluation ('M&E') systems are, in principle, designed to monitor the flow of project inputs and outputs and to evaluate not only the effects of intervention in a restricted way but also the impact in a wider sense.<sup>40</sup> 'M&E' systems usually have two laudable objectives: to place more emphasis on the need for systematic monitoring and routine process evaluation, and to keep the total costs of monitoring and evaluation low. However, they are too often depicted as a kind of 'miracle' package of process evaluation and impact evaluation, together with involvement of local personnel which, on top of all this, would also be low in costs.<sup>41</sup>

<sup>36</sup>A. Thomson, 'Food aid as income transfer', in M.J. Forman, ed, *Nutritional Aspects of Food Aid*, ACC/SCN (FAO), United Nations, Rome, 1986, pp A38-A46.

<sup>37</sup>J. Katona-Apte, 'Women and food aid: a development perspective', *Food Policy*, Vol 11, 1986, pp 216-222.

<sup>38</sup>Burgess, *op cit*, Ref 2.

<sup>39</sup>Mention should also be made of nutrition surveillance, although it falls outside the scope of this review. In its general sense nutritional surveillance systems involve data collection for national planning purposes, with data being drawn mostly from already existing sources. The systems include widely differing areas of information such as weather conditions, agricultural production projections, food prices and the nutritional status of the population. Routine monitoring of the nutritional state of young children is a part of nutritional surveillance, and is sometimes referred to as nutritional screening. J.P. Habicht and J. Mason, 'Nutrition surveillance: principles and practice', in D.S. McLaren, ed, *Nutrition in the Community*, Wiley, New York, 1983, pp 217-244; and J.B. Mason, J.P. Habicht, H. Tabatabai and V. Valverde, *Nutritional Surveillance*, World Health Organization, Geneva, 1984.

<sup>40</sup>R.I. Miller and D.E. Sahn, 'Built-in evaluation systems for supplementary feeding programmes - why and how', in D.E. Sahn, R. Lockwood and N. Scrimshaw, eds, *Methods for the Evaluation of the Impact of Food and Nutrition Programmes*, United Nations University, Tokyo, 1984, pp 265-287.

<sup>41</sup>H. Gotzman, 'Improved project design, monitoring and evaluation (some recent WFP experience)', in *UNICEF/WFP Workshop on Food Aid and the Well-Being of Children in the Developing World*, UNICEF, New York, 1986, pp 143-154.



The proposed 'M&E' systems have severe limitations with regard to the topics that can be covered and subsequent insights that can be gained, but they also face severe restrictions on the type and reliability of data that can be collected, and the availability, accuracy and motivation of project staff. The few examples of such systems functioning in actual practice show that they only exist by the grace of qualified research personnel and that they are nevertheless restricted in scope.<sup>42</sup> Impact evaluation requires data to be aggregated and compared across carefully selected subgroups, while often data on non-participants are also needed for purposes of comparison. Moreover, information is often needed that cannot be collected routinely by project staff, either because they lack the necessary training or simply because they lack the time. There is no doubt that 'M&E' systems require so-called extra or special studies which either address typical impact issues or test behavioural assumptions underlying the intervention.<sup>43</sup> In fact, in this connection, it may be more useful to speak of different stages of evaluation.<sup>44</sup>

Ultimately the aim of monitoring and evaluation is to gain information about the daily functioning of nutrition intervention programmes and about their effects. The quality of such information and possible ensuing insights are determined by the issues that are addressed. The more fundamental these issues are, the higher the costs of research. In this sense, it is inevitable that as soon as 'M&E' systems are genuinely directed at impact evaluation, they will run into the same problems that others have come to experience. In its final form impact evaluation is concerned with behavioural research, and this is expensive, requires expert personnel and takes time.

## Conclusion

Experiences with rural development in developing countries have shown that it is generally unrealistic to expect spectacular improvements in economic conditions or well being as a result of one particular intervention. Food and nutrition form an essential part of daily life and are closely interwoven with many other aspects of human existence and social organization, and these do not change overnight. Consequently, there should not be overblown expectations of the impact of any particular kind of intervention, and this in turn influences what can be demanded and expected of programme evaluation. Targeting of interventions is a *sine qua non*. In this way scarce resources can be better utilized and – perhaps equally important – the minor influence of interventions can be capitalized so that impact becomes more pronounced, thus measurable, subsequently researchable, and finally improvable.

The essential features of evaluations under these conditions are threefold. Firstly, there are the many methodological difficulties which almost invariably lead to imperfections of design, which either weaken the power of the evaluation, or restrict its scope. Secondly, the effects of most nutrition interventions are at best modest. Finally, nutrition programmes in tropical countries have to operate in a great variety of cultural settings, which not only raises practical implementation problems but which must inevitably also have a strong bearing on the interpretation of evaluation findings.

The question must be raised whether impact evaluation has not often

<sup>42</sup>Eg, S.J. Munger, 'The generalizability of evaluation protocols for nutrition education', in B. Schurch, ed, *Evaluation of Nutrition Education in Third World Communities*, Hans Huber, Bern, 1983, pp 183–226.

<sup>43</sup>Sahn, *op cit*, Ref 30, pp 1–14

<sup>44</sup>Mason and Habicht, *op cit*, Ref 35.



received undue emphasis, to the detriment of process evaluation. One must seriously doubt the need and feasibility for inbuilt, that is continuous, evaluation of impact among recipients if the impact is of such an elusive nature as behavioural changes among mothers and developmental improvements of children. Process evaluation, on the other hand, is necessary and there is a day-to-day need for it, so as to assure the proper delivery of services. In principle, the necessary monitoring can be done routinely by permanent project staff, at the same time providing an essential data source on which to graft later impact evaluation.

Evaluations, like the actual interventions, have to operate under limiting conditions which restrict their potential contributions, and that certainly does not make them the final arbiters that they are often expected to be. Evaluation is a tool, a tool that has to be used carefully and skilfully and as such is an important, but modest, means to achieve improvements and to realize the necessary targeting. H.W. Riecken<sup>45</sup> has remarked that one can only expect the quality of evaluation for which one is prepared to pay, financially, but also in terms of effort and interference with daily operations. Programme managers, officials and politicians must weigh the costs against the possible insights that can be gained. Ultimately, the costs and benefits of evaluation itself have to be weighed against each other.

<sup>45</sup>H.W. Riecken, 'Practice and problems of evaluation: a conference synthesis', in R.E. Klein, M.S. Read, H.W. Riecken, J.A. Brown, A. Pradilla and C.H. Daza, eds, *Evaluating the Impact of Nutrition and Health Programs*, Plenum Press, New York, 1979, pp 363-386.

Monitor is a quarterly review of selected documents emanating from the four international agencies based in Rome – the Food and Agriculture Organization of the UN (FAO), the UN/FAO World Food Programme (WFP), the World Food Council (WFC) and the International Fund for Agricultural Development (IFAD). Compiled by Ram Saran, this issue's Monitor looks at agricultural price policies.

## Design and implementation of agricultural price policies

### *Agricultural Price Policies: Issues and Proposals*

Food and Agriculture Organization of the United Nations, FAO Economic and Social Development Series No 42, Rome 1987

### *Evaluation of Food Aid for Price Stabilization and Emergency Food Reserve Stocks: a Study of WFP-assisted Projects in Tanzania, Botswana, Mauritania, Niger and Mali in 1984*

UN/FAO World Food Programme, Occasional Paper No 2, Rome, 1985

Agricultural price intervention is widely used as an instrument to influence agricultural prices. Through such interventions, most governments seek to create a positive price climate for agricultural growth. Another objective is to ensure reasonable prices for consumers and manufacturers of food commodities and agricultural raw materials. This instrument is also used to influence farm incomes, to regulate exports or imports, or to provide revenue to governments. The importance of each of these objectives and the measures used to attain them differ from country to country depending upon such factors as: the importance of the agricultural sector in the country's economy; the proportion of agricultural production that is marketed; the stage of agricultural development; the surplus/deficit of agricultural commodities produced in the country; the extent to which non-price factors (eg technological factors) can be relied upon to expand agricultural output; and the feasibility of income transfers from the non-agricultural to the agricultural sector.

The dominance of the agricultural sector in the economies of the developing countries, together with the gravity of the food and agricultural situation which many of them

face, has caused concern among governments and international organizations that the price policies pursued by these countries are perhaps inappropriate and need modification. The rising costs of agricultural pricing policies followed by developed countries have also led to consideration of ways to modify or eliminate these policies. Ways of improving price incentives are also being given serious thought in some centrally-planned countries. The FAO study on agricultural price policies provides an insight into price developments and problems of designing and implementing price policies in countries of all economic classes – developed, developing and centrally-planned. Its focus, however, is on developing countries.

The FAO study is the result of considerable analytical work and experience of those engaged in formulating, applying and evaluating agricultural price policies. Preparation of this study involved contributions from country studies by FAO and other organizations, a quantitative analysis based largely on a set of data for 38 developing countries, discussions at FAO intergovernmental meetings, an extensive review of professional literature and a discussion of this review by an expert panel, and reports of consultants and substantive comments by FAO staff with field experience in the subject.

The FAO report begins with a summary of price developments and policy changes which have been introduced by national authorities in recent years. The FAO analysis shows that a rise in the price of food staples will reduce the calorie consumption of the poor by about the same proportion. In other words, an increase of, say, 10% in price would cut down their food intake by 10% or by about 200 calories, which would have a serious effect on health, energy and resistance to disease. The elasticity of agricultural output has been computed at 0.2–0.5, which suggests that the total production response to price movements is 'large enough for prices to matter but low enough to rule out the possibility of increased price incentives alone fuelling sustained growth in agriculture'. This is because agricultural production depends not only on prices but also on the adequacy of agricultural infrastructure which encompasses suitable technologies, input and credit supply, and public investments in such areas as irrigation, research, extension, marketing and

Mr Ram Saran is a former head of the Food Security and Food Aid Policy Group in the Commodities and Trade Division of the FAO. The views expressed here are his own and do not necessarily represent any of the organizations concerned.